SUPPLEMENTAL APPENDIX for
"Improved Density Forecasts Based on Mixed Frequency Data: A Focused Bayesian Approach"

A: Details on the benchmark models

The key benchmark models are the plain vanilla GARCH model, and the GARCH-RV model. As described in Section 2, both benchmarks are nested within the HF-GARCH specification. Specifically, the GARCH(1,1) conditional volatility equation takes the form of

$$\sigma_t^2 = \beta_0 + \alpha_1 y_{t-1}^2 + \beta_1 \sigma_{t-1}^2,$$

and the GARCH-RV(1,1) specification is

$$\sigma_t^2 = \beta_0 + \alpha_1 \sum_{i=1}^{M} y_{(t-1)_i}^2 + \beta_1 \sigma_{t-1}^2.$$

In the predictive analysis conducted in 4, we employ the Gaussian innovation terms, and apply the adaptive MH algorithm to estimate the parameters $(\mu, \beta_0, \alpha_1, \beta_1)^\top$ for both the likelihood-based and focused predictions. As in the case of the HF-GARCH, we target the MH acceptance ratio between 30-70%. All resulting predictive distribution integrates out all model unknowns, as per equation (9) in the main text.

B: Comparison between Bayesian and MLE predictions

We compare our Bayesian predictive distributions to those produced by frequentist MLE. Without the toolkit of the hierarchical prior for regularization, we apply Ridge-type regularization as in Eilers and Marx (1996) as penalty term to the likelihood function. As such, the MLE predictive, $p_{MLE}(y_{T+1}|\mathbf{y}_{1:T}, \eta_{1:T}, \hat{\varphi}_{MLE})$, is constructed from the MLE estimate

$$\hat{\boldsymbol{\varphi}}_{MLE} = \arg\max_{\boldsymbol{\varphi}} \left(\log p(\boldsymbol{y}_{1:T}|\boldsymbol{\varphi}, \boldsymbol{\eta}_{1:T}) - \lambda_T ||\Delta\boldsymbol{\Gamma}|| \right),$$

where $||\Delta\boldsymbol{\Gamma}||$ denote the L_2 norm of the first difference in the parameter function $\boldsymbol{\Gamma}$ and the penalty $\lambda_T = \tilde{\lambda}T$ varies with the sample size to accommodate the scale of the likelihood. We consider the range $\tilde{\lambda} \in [0, 1]$, with the optimal value of $\tilde{\lambda}$ dictating this penalty chosen from predictive cross validation. The value of $\tilde{\lambda} = 0$ indicates zero penalty, with the parameter function $\boldsymbol{\Gamma}$ allowed to have full flexibility; while the value $\tilde{\lambda} = 1$ regularizes $\boldsymbol{\Gamma}$ to a flat line, coinciding the estimated model to the GARCH-RV special case.

Table A1 reports the predictive log score for each of the thirteen indices produced by the Bayesian posterior and the MLE estimate described above. The log scores are computed over the out-of-sample period investigated in Section 4. The table reveals that the MLE and Bayesian predictive distributions produce similar overall accuracy, as measured by the log score. For the majority of the indices, the Bayesian predictive produces larger log score. We performed the Hansen (2005) Superior Predictive Ability (SPA) test to assess whether the predictive that generate larger log score indeed perform better statistically. Out of the thirteen indices, the Bayesian prediction is considered statistically superior to the MLE prediction in seven cases. The MLE prediction dominates only in one case, and the two methods perform equivalently in the remaining cases.

C: Predictive performance with longer lag order

We further compare the predictive performance of the HF-GARCH(p,q) model to assess the impact of the choice of lag p. As in the analysis of the main paper, we compare two choices: $p = 1$, implying the inclusion of intraday data only in the day leading up to the prediction period; and $p = 21$, implying the inclusion of intraday data of the month leading up to the

Table A1: Comparison of the MLE and Bayesian predictive log score of the HF-GARCH(1,1) model with $B = 5$. Bolded values represent the larger log score value. Statistically superior predictive according to the SPA test are denoted by * and ** for the 10% and 5% statistical signficance, respectively.

	MLE	Bayes		MLE	Bayes
SPY	-1.096	**-1.069****	SPXEW30	-0.955	**-0.954**
MID	**-1.287****	-1.292	SPXEW35	-1.156	**-1.154**
SPCY	**-1.385**	-1.386	SPXEW40	-1.227	-1.227
SPXEW10	-1.696	**-1.695***	SPXEW45	-1.259	**-1.257****
SPXEW15	-1.262	**-1.259****	SPXEW50	-1.328	**-1.313***
SPXEW20	-1.156	**-1.155****	SPXEW55	-1.222	**-1.220**
SPXEW25	-1.213	**-1.210***			

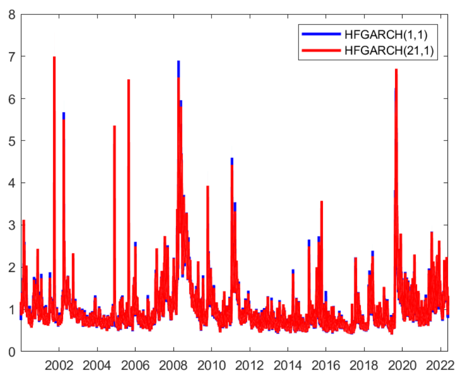
prediction period. For this predictive exercise, we estimate the models using $B = 5$ for the case of $p = 1$ and $B = 15$ for the case of $p = 21$ via MLE with regularization penalty as described above. The predictive log score are documented in Table A2. Here we observe that while the larger model with $p = 21$ produces larger log score over the out-of-sample prediction period for nine of the thirteen series, the larger model is not statistically superior in its predictive ability in all but one of these nine indices.

Table A2: Comparison of the predictive log score of the HF-GARCH(p,q) model for the case of $p = 1$ with $B = 5$ and $p = 21$ with $B = 15$. Both models are estimated using MLE. Bolded values represent the larger log score value. Statistically superior predictive according to the SPA test are denoted by * and ** for the 10% and 5% statistical signficance, respectively.

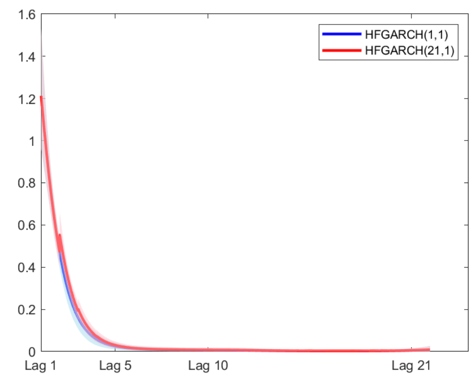
	$p = 1$	$p = 21$		$p = 1$	$p = 21$
SPY	**-1.096****	-1.107	SPXEW30	-0.955	**-0.953**
MID	**-1.287**	-1.288	SPXEW35	-1.156	**-1.154**
SPCY	-1.385	**-1.384**	SPXEW40	-1.227	**-1.226**
SPXEW10	-1.696	**-1.694**	SPXEW45	**-1.259**	-1.259
SPXEW15	-1.262	**-1.258**	SPXEW50	**-1.328**	-1.331
SPXEW20	-1.156	**-1.154**	SPXEW55	-1.222	**-1.220**
SPXEW25	-1.213	**-1.210****			

Additional comparison of the fitted volatility and implied intraday weights from the HF-GARCH(1,1) and HF-GARCH(21,1) for the MID and SPCY indices are provided in Figure A1.

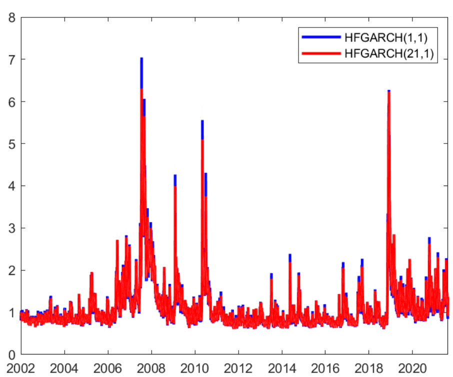
Panel a): MID Fitted σ_t^2



Panel b): MID Model implied intraday weights



Panel c): SPCY Fitted σ_t^2



Panel d): SPCY Model implied intraday weights

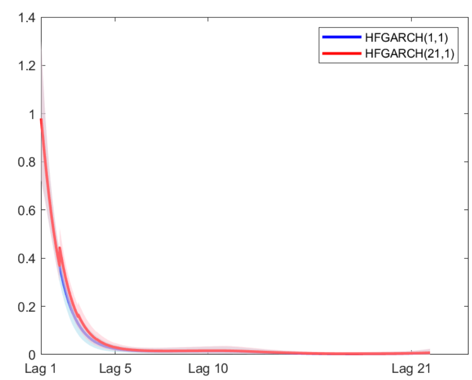


Figure A1: Comparison of the the HF-GARCH(1,1) and HF-GARCH(21,1) models for the MID and SPCY indices.

D: Supplementary empirical results

Further results on the intraday weights implied by the focused predictions constructed from the HF-GARCH(1,1) for the indices not presented in the main paper are presented in Figures A2 and A3. The representative predictive densities produced for the median realized volatility day for the remaining indices not presented in the main paper are given in Figures A4 and A5.

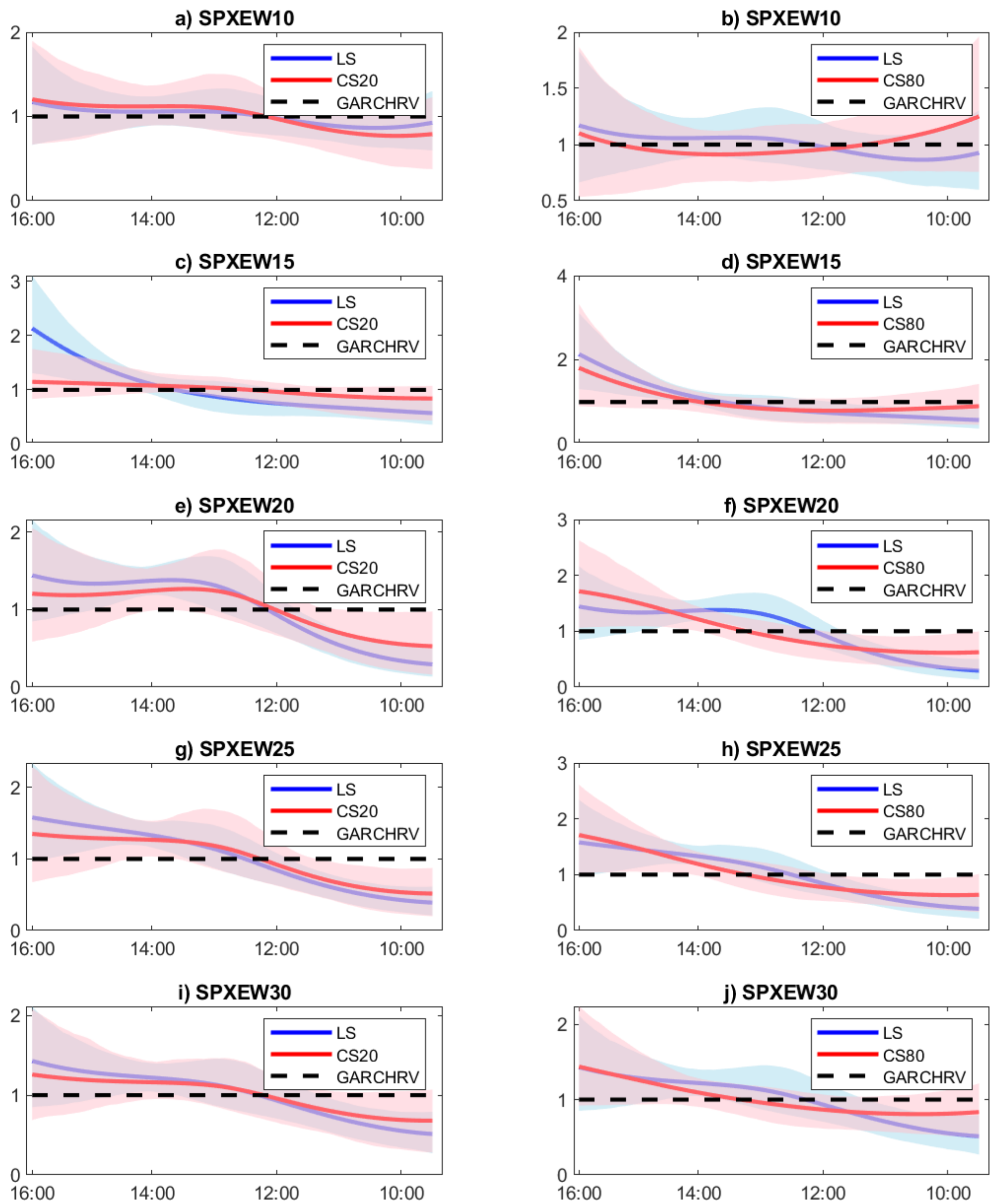


Figure A2: Normalized intraday weights for the industrial indices: SPXEW10, SPXEW15, SPXEW20, SPXEW25 and SPXEW30.

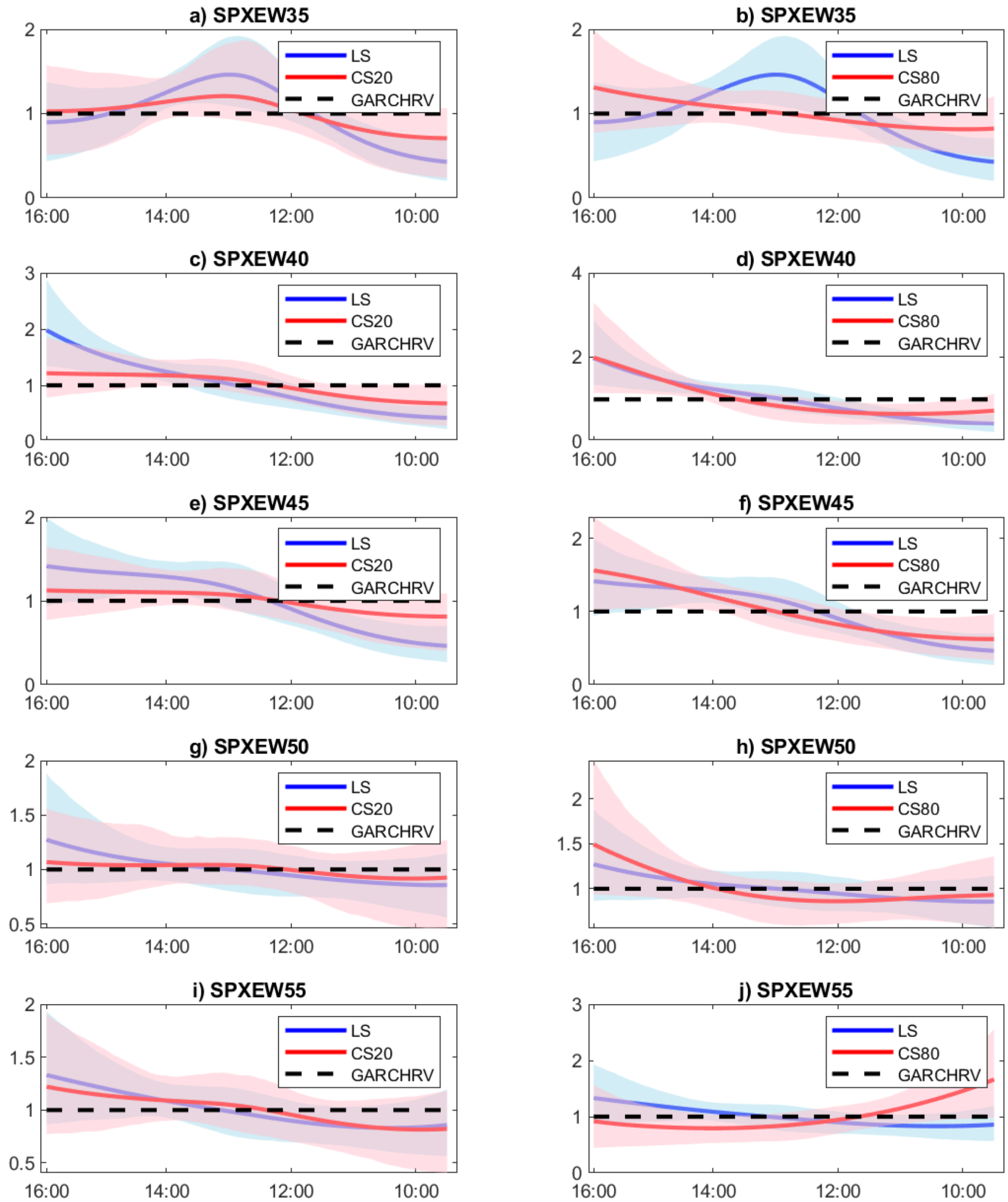


Figure A3: Normalized intraday weights for the industrial indices: SPXEW35, SPXEW40, SPXEW45, SPXEW50 and SPXEW55

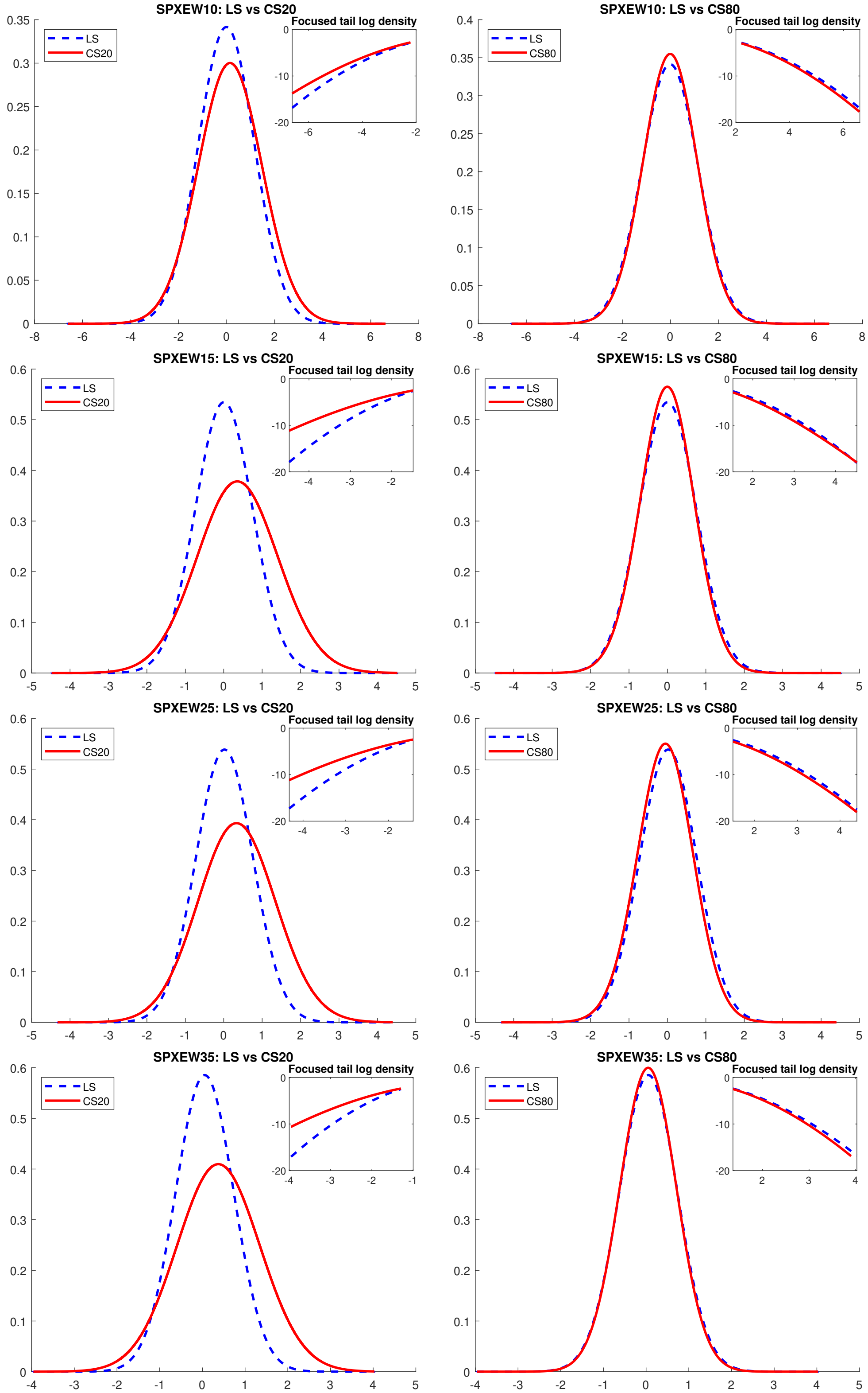


Figure A4: Representative predictive densities for the SPXEW10, SPXEW15, SPXEW25 and SPXEW35 indices on a day where the realized volatility is at the respective median values.

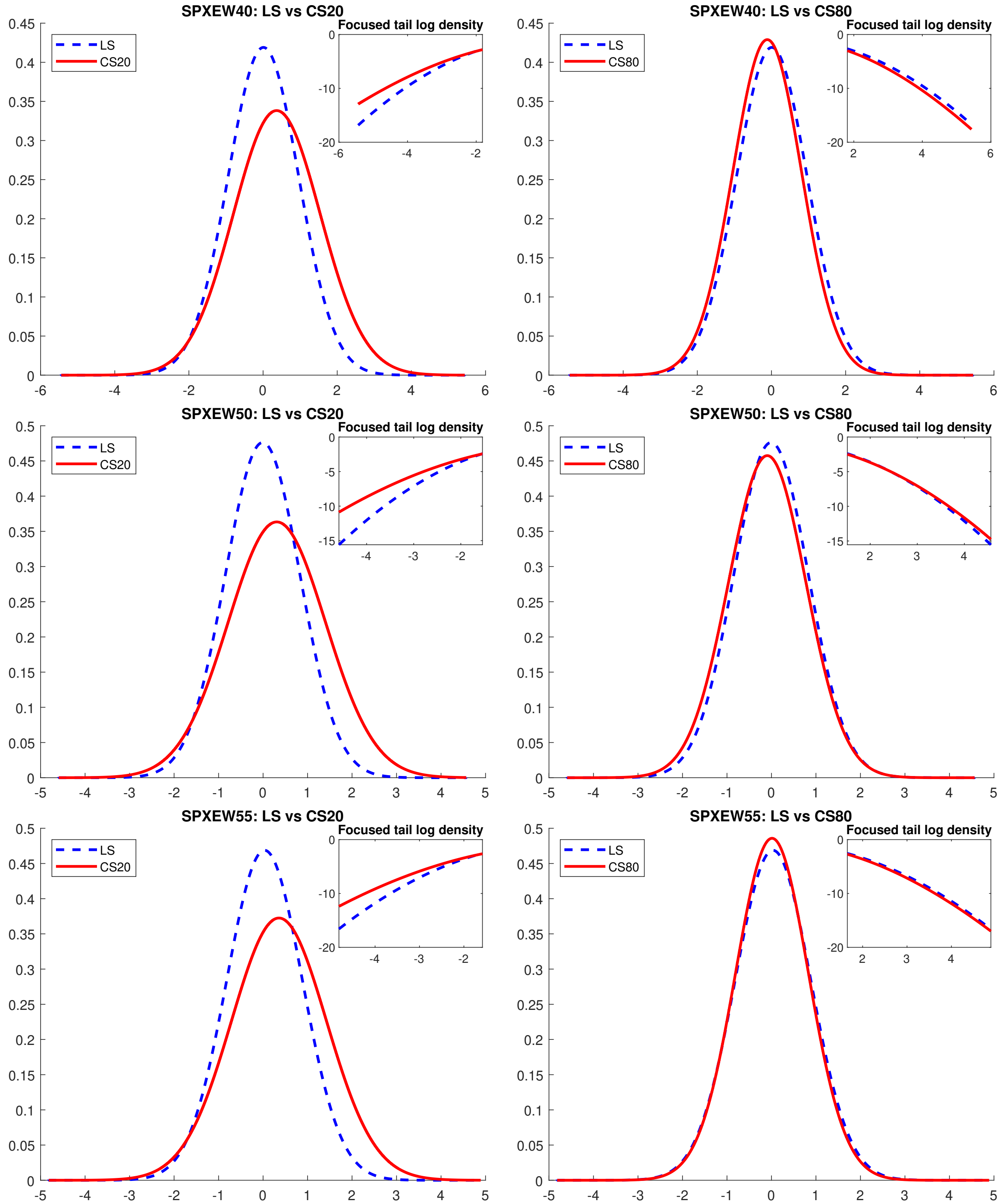


Figure A5: Representative predictive densities for the SPXEW40, SPXEW50 and SPXEW55 indices on a day where the realized volatility is at the respective median values.